new media & society

London, Thousand Oaks, CA and New Delhi
Vol8(5):825–842 [DOI: 10.1177/1461444806067737]

ARTICLE

The mp3 as cultural artifact

JONATHAN STERNE
McGill University, Canada

Abstract

The mp3 lies at the center of important debates around intellectual property and file-sharing, but it is also a cultural artifact in its own right. This article examines the design of the mp3 from both industrial and psychoacoustic perspectives to explain better why mp3s are so easy to exchange and the auditory dimensions of that process of exchange. As a container technology for recorded sound, the mp3 shows that the quality of 'portability' is central to the history of auditory representation. As a psychoacoustic technology that literally plays its listeners, the mp3 shows that digital audio culture works according to logics somewhat distinct from digital visual culture.

Key words

digital audio • digital format • file-sharing • listening • mp3 • psychoacoustics • sound • sound recording • technology and culture • visual culture

For the last seven years or so, the mp3 has occupied center stage in the world of digital audio formats. It has been the subject of academic papers, court cases, congressional and parliamentary hearings and countless magazine and newspaper articles. Mp3 trading has been the case in point in a major international controversy over the status of intellectual property, copyright and the economics of entertainment. A whole series of authors have argued that the debate over intellectual property is incredibly important for intellectuals, academics, artists and anyone else who works with ideas (see e.g. Bettig, 1997; Burkart and McCourt, 2004; Jones, 2000; Lessig, 2000, 2002; McCourt and Burkart, 2003; McLeod, 2001, 2005). Writings on

mp3s and file-sharing almost uniformly sound a note of crisis, as if the battle over mp3s and intellectual property is the most important cultural conflict of our time.

Therefore, it is surprising how little of the common sense of technology studies has been applied to mp3s. Scholars in a range of fields – philosophy of technology, science and technology studies and the cultural study of technology – have all advocated the study of technologies as artifacts. Philosopher Langdon Winner writes that technological artifacts 'embody specific forms of power and authority' (1986: 19). He groups the politics of technologies into two main categories: 'instances in which the invention, design, or arrangement of a specific technical device or system becomes a way of settling an issue in the affairs of a political community', and '"inherently political technologies", man-made systems that appear to require or be strongly compatible with particular kinds of political relationships' (1986: 22). In Winner's heuristic, the mp3 partakes of both categories: it originated as an attempt to solve the problem of exchangeable formats across segments of the media industry and it may require particular social and cultural systems of both intellectual property and listening.

The mp3 is an artifact in another sense. The mp3 is a crystallized set of social and material relations. It is an item that 'works for' and is 'worked on' by a host of people, ideologies, technologies and other social and material elements. Writers in the social construction of technology and actor–network theory traditions (e.g. Bijker, 1995; Latour, 1996; Pinch and Bijker, 1984) have focused on the relation of human and non-human actors in the construction of technologies, showing how technologies come together from what one might consider otherwise as disparate elements. Cultural studies of technology have been more concerned with broader accounts of social context and stratified social power as they shape technologies and as technologies are implicated in these contexts (see e.g. Slack, 1984; Stabile, 1994; Wise, 1997). But all these approaches point to the artifactual nature of technologies such as the mp3. They urge us to consider it as a result of social and technical processes, rather than as outside them somehow. Uncovering that process is not simply a matter of showing the artificiality or 'constructedness' of the mp3, although that is part of the project. This article will use the mp3 as a tour guide for social, physical, psychological and ideological phenomena of which otherwise we might not have been fully aware. It will consider the mp3 as an artifact shaped by several electronics industries, the recoding industry and actual and idealized practices of listening.

Of course, this is not the first cultural study of mp3s. Kembrew McLeod (2005) notes that because the mp3 format is software, its uses are somewhat less determined than hardware and, even there, uses can change. Steve Jones (2002) has argued cleverly that the mp3 is an occasion to bring questions of

distribution to the forefront of cultural studies. Yet in most accounts, writers still represent the mp3 itself as a mute, inert object that 'impacts' an industry, a social environment or a legal system. The writing on the subject most often takes the form of the mp3 as either 'given' or obvious, with little further thought on the matter required for addressing real legal and economic issues. At the same time, surprisingly little discussion has occurred around the aesthetic dimensions of mp3s, whether by that one means the experience of mp3 listening, the sound of mp3s themselves or the meanings that the form of the mp3 might take. Discussions of the sound of mp3s have been limited largely to audio engineers and audiophiles, who range from dismissals on the basis that mp3s sound 'bad' (e.g. Atkinson, 1999) to analyses of the sonic limitations of mp3s as a 'problem' (e.g. Eide, 2001). In the academic world, one could read for a long time before confronting the fact that mp3s are sound files. Yet to note an absence is not enough. After all, if the substantial dimensions of 'the mp3 question' are law and economics, one might reasonably assume that concerns about the mp3 'as technological form' and sonic object take a back seat.

This article advances an alternate position. A robust understanding of the technological and aesthetic dimensions of the mp3 provides an important context for discussions of the legal, political economic and broader cultural dimensions of file-sharing. By examining the mp3 as an auditory technology, it reveals some important dimensions of the relationship between the so-called 'new' media[1] and the human body that have been neglected largely by scholars who privilege the visual dimensions of new media. In short, it will show that a gestural, tactile form of embodiment is the requirement and result of digital audio. This contrasts greatly with the mentalist and self-conscious disembodiment that some scholars still describe as a central feature of virtual space.

To borrow a term from Lewis Mumford, the mp3 is a 'container technology'. Mumford wrote that technology scholars' emphasis on tools over containers 'overlooks' their equally vital role (see Mumford, 1959, 1966). He postulated that one reason why container technologies are neglected often in the history and philosophy of technology is that usually, they are coded as feminine. While gender coding may be a bit dated, Mumford did have a point about activity and passivity, which are still often gender coded. More recently, feminist scholar Zoë Sofia (2000) has picked up Mumford's thread. While she qualifies Mumford's argument – that container technologies may be as connected with men as with women – they are still metaphorized often as feminine. But Sofia argues that the misogyny story is only part of the explanation for the neglect of container technologies: 'to keep utensils, apparatus and utilities[2] in mind is difficult because these kinds of technological objects are designed to be unobtrusive and . . . make their presence felt, but not noticed' (2000: 188). Indeed, this

is the mode in which mp3s work: they are important precisely because they are useful but do not call attention to themselves in practice. They take up less space than other kinds of digital recordings and when they are listened to, they are experienced as music, not as file formats. Thus, we should not be surprised to find media among Sofia's many examples of container technologies.

The rest of this article explores the mp3 as a container technology for sound recordings. Mumford and Sofia both use the term 'apparatus' to describe a container that transforms as it holds (Sofia, 2000). The mp3 clearly belongs to this category, but it differs in one important way: it is a container for containers. Like an oven that holds a casserole and transforms its contents, the mp3 is a holder for sound recordings. It is a media technology designed to make use of other media technologies. As we will see below, the transformations effected by mp3 encoding are themselves heavily-directed cultural practices. Mp3s contain within them a whole philosophy of audition and a praxeology of listening. As a philosophy of audition, the mp3 makes use of the limitations of healthy human hearing. One might even say that the mp3 is a celebration of the limits of auditory perception. The anticipated praxeology of listening encoded in each mp3 emphasizes distraction over attention and exchange over use. If that were not enough, the technology itself is perfectly and lovingly shaped for the very purposes to which it is not supposed to be put: the mp3 is perfectly designed for illegal file-sharing. Thus, the next section explains the form of the mp3 and what makes it different from other kinds of recording technologies and then offers an analysis of the mp3 as a cultural artifact.

THE SUPPLY OF MP3s

The point of mp3s is to make audio files smaller through data compression[3] so that they are easier to exchange in a limited bandwidth environment such as the internet, and easier to store in a limited dataspace environment, such as a hard drive. This section discusses the reasons why a consortium of communication industries built the mp3 to be so portable. Then it explains the psychoacoustic dimensions of mp3s, which are the crucial technical and cultural components of their portability. In short, the mp3 was designed by an electronics industry interested in maximum compatibility across platforms, which would allow for easy exchange of files. At the same time, the mp3 uses a specific form of data compression based on a model of how the human ear works. Therefore it is a machine designed to anticipate how its listeners perceive music and to perceive for them. Both explanations are a bit 'gear-headed', but like the 'detour through theory' so central to good cultural studies writing (Hall, 1992: 283), this 'detour through technology' will help us to reconstitute our object of study and thereby gain new insights into hitherto hidden auditory dimensions of digital media culture.

MP3 stands for MPEG-1, Layer-3. MPEG (Motion Picture Experts Group) is a consortium of engineers and others formed with the support of the International Standards Organization (ISO) and the International Electrotechnical Commission.[4] MPEG started out in 1988 as an ad hoc group that aimed to standardize data compression schemes in the broadcast, telecommunications and consumer electronics industries. In collaboration with academics, all of the big corporate players had a role and an interest, even as internal divisions within corporations sometimes led them to create proprietary standards while supporting research on standardized compression. For those who believed in standardization, one format was needed that could cut across digital technologies as diverse as compact discs, digital video, high-definition television, teleconferencing and satellite communications, to name just a few of the disparate industries interested in a shared standard (Mitchell et al., 1997).[5] The MPEG-1 standard is broken down into five 'parts': part 1 deals with systems aspects; part 2 with video; part 3 with audio compression; part 4 with compliance testing; and part 5 with software implementation (Mitchell et al., 1997). But the mp3 does not stand for the third part, it stands for the third layer of audio coding in the MPEG standard. So, mp3 is really layer 3 of part 3 of the MPEG standard. (This level of detail is offered to make a point: the audio dimension of the MPEG standard was, at the outset, a very local concern – a piece of a larger project which aimed to standardize compression across all forms of digital media.)

To borrow another phrase from Zoë Sofia, the MPEG standard carried with it a certain 'logic of resourcing and supply' (2000: 195–6). The logic was this: once standardized, data could be moved with ease and grace across many different kinds of systems and over great distances frequently and with little effort. This was the dream of the Motion Picture Experts Group and its many benefactors. If one is looking for the cultural origins of the promiscuity among illegal file-sharers, one need look no further than this founding moment. The possibility for quick and easy transfers, anonymous relations between provider and receiver, cross-platform compatibility, stockpiling and easy storage and access – all were built into the MPEG form itself long before the age of Napster, Gnutella, Hotline, iTunes and Rio.

Attempts to commercialize the mp3 and its likely successor, Advanced Audio Coding (AAC), make use of various digital rights management algorithms that make it more difficult to share files. For instance, Apple's iTunes Music Store and RealNetworks use incompatible technologies to prevent sharing. These recent innovations do mark a recognition on the part of the industry that the compatibility of the mp3 is part of the reason why it is so widely shared. But the net result of such practices is to require users who legally purchase mp3s from iTunes and RealNetworks to use two separate programs for playback. As Patrick Burkart and Tom McCourt

(2004) argue, digital rights management is hampered by vast incompatibilities across software and hardware and an industry climate hostile to a shared standard. In other words, it is probably easier to install a Gnutella client and acquire illegal files than it is to manage two separate mp3 collections and two separate playback programs.

Through its design as a portable container technology, the mp3 has been ascribed the status of a thing in everyday practice, even though it is nothing more than a format for encoding digital data. Both listeners and companies that sell mp3s (or the equipment to play them) readily talk of mp3 collections analogously to record collections or book collections. Articles on the original incarnation of Napster, for example, tended to describe it as a 'program for searching other people's mp3 collections' (e.g. McCollum, 2000; Wood and Jenish, 2000). More recently, magazine advertisements for Apple's iPod and iPod Mini extol their virtues by listing the number of 'songs' that each device can hold (Apple, 2004). Further, mp3s have been objectified as articles of intellectual property in the US legal system, Canadian legislature and in several other countries as well (Evangelista, 2003; Krim, 2003; Lazin, 2003).

This raises the problem of use-value and exchange-value. Use-value, first identified by John Locke (1692) and elaborated by Karl Marx and others, is a perspective on value which treats commodities in terms of their actual utility. Exchange-value, meanwhile, is the market value of a commodity. While this would appear to be an arbitrary relationship, Marx argues – following Adam Smith – that exchange-value is actually based on the labor required to make the commodity, once that labor is exchanged for money which in turn can be traded for the commodity (Marx, 1967[1867]; Smith, 1993[1776]). Thus, we have a bifurcation of value: use-value, which is about the work of expenditure; and exchange-value, which is about the work of creation. Most interesting for our purposes is the relationship of exchange-value to music. In his romantic and polemical book on music, Jacques Attali argues that sound recording occasions a shift from use-value to exchange-value in music:

> We must not forget that music remains a very unique commodity; to take on meaning, it requires an incompressible lapse of time, that of its own duration. Thus the gramophone, conceived as a recorder to stockpile time, became instead its principle user. Conceived as a word preserver, it became a sound diffuser. The major contradiction of repetition is evident here: *people must devote their time to producing the means to buy recordings of other people's time*, losing in the process not only the use of their time, but also the time required to use other people's time. Stockpiling thus becomes a substitute, not a preliminary condition, for use. *People buy more records than they can listen to. They stockpile what they want to find the time to hear.* Use-time and exchange-time destroy one another. (Attali, 1985: 101; emphasis in original)

One can see the problem immediately with mp3s in Attali's formulation: usually, people do not buy them. Attali argues that recording substitutes exchange-value for use-value because people do not have the time to listen to all of the recordings that they may buy. Yet compared to the number of mp3s freely given and received through file-sharing, few meet the basic definition of exchange-value: they are not paid for and therefore do not require as much labor (in exchange for a wage or salary) to procure. Further, the exchange itself does not deprive the original owner of the file's use. The peculiar status of the mp3 as a valued cultural object which can circulate outside the channels of the value economy is one of the fundamental, enabling conditions for the intellectual property debates that surround it. Epochal proclamations are tempting when confronted with this state of affairs: one could say that if recording shifted music from use-value to exchange-value, then digitization in the form of the mp3 liberates recorded music from the economics of value by enabling its free, easy and large-scale exchange.

In use, mp3s can seem a bit like mollusks without their shells – recorded music without the commodity form – since generally they are not exchanged for money. One response to this condition would be to consider the internet as a gift economy, as Richard Barbrook (1998) has done so elegantly and provocatively. But if the mp3s are mollusks free of their shells, they still need air and water: listeners must still pay for the descendents of the gramophone and the record dealer: computers, speakers, internet connections (or membership in institutions such as universities that provide access to such things) and possibly other playback devices such as Rios or iPods. Further, most of the recordings now available in mp3 form once lived in a money economy, paid for by record companies (or less often, independent musicians) who in turn put them up for sale in the hope of realizing a profit. Thus, whatever side we take in the debates over intellectual property and digital rights management, the question of value persists. If for no other reason, we know this because users continue to desire, collect, stockpile – and yes, use – mp3s.

'For a collector,' wrote Walter Benjamin, 'ownership is the most intimate relationship that one can have to objects' (1968[1936]: 67). That one can collect mp3s suggests that they appear to users as cultural objects, even if they are not, in any conventional sense, physical objects that can be held in a person's hand. In a review of Traktor, a software DJ program, Philip Sherburne pauses to note that digital audio formats and their manipulation on computers reflect

> the ongoing dematerialization of music (or perhaps a better term would be 'micromaterialization' since even mp3s live in silicon, invisible as they may seem). More and more, our collections exist not on our record shelves, but in our iPods and hard drives. (2003: 46)

This is key: although mp3s exist as software, people tend to treat them like objects (and indeed, the argument here is that we should analyze them as artifacts) perhaps because they are used to handling recordings as physical things. But because of their micromaterialization, users can handle mp3s quite differently from the recordings they possess in a more obviously 'physical' form such as a record or compact disc (CD), even though they may talk about mp3s as if they are physical objects. Consider this review of the iPod and iTunes Music Store:

> [Y]ou'll even find that you listen to music in new ways. Recently the Talking Heads' sublime 'Heaven' popped up on my jukebox in random play mode; I'd owned the CD for years but hadn't played it much and never noticed this amazing song. That kind of discovery happens all the time now that my music collection has been liberated from shiny plastic disks. (Regnier, 2003: 113)

There are really two kinds of object indicated in the quote: there are objects that can be collected, which include mp3s, and there are objects that can be touched (in some conventional sense), such as CDs but unlike mp3s. All this is to say that if we accept the language of the materialization and dematerialization of music,[6] mp3s present us with an interesting bifurcation. Users refer to the dematerialization of music in discussing their practices of use, but they insist on treating music as a cultural object when they discuss their possession of the music.

INSIDE THE MP3

Because it is so small, the mp3 format makes collecting all that much easier: an entire collection can fit in a relatively small space. An mp3 takes an existing CD-quality digital audio file and removes as much data content as possible, relying on listeners' bodies and brains to make up the difference. For example, a three-minute stereo CD file takes up about 30 megabytes of disk space; a three-minute mp3 of average quality takes up 3 to 4 megabytes of disk space. This is accomplished through a variety of filters and processes. The patents for the audio section of the MPEG standard continue to be held mostly by a German company named Fraunhofer IIS. A traditional zip utility gets rid of redundant data to make a file a little smaller. However, if you were to zip a CD track, it would not get much smaller. Fraunhofer's basic innovation was to use a mathematical model of human auditory perception to allow for greater data compression in mp3 files. In essence, the file is designed to figure out what you will not hear anyway and to get rid of the data for that portion of the sound.

Although it is a data file, it has been suggested already here that users treat mp3s as cultural objects. Mp3s are like other technologies in another important way: they are assembled by other technologies. The name for a program that assembles mp3s is an encoder. For the purposes of this

argument, an encoder will be treated like any other container technology that transforms its contents. The encoder takes an existing digital recording and processes it through six related steps (the following discussion is based on Hacker, 2000).

(1) The mp3 encoder breaks the signal down into smaller pieces called 'frames', each lasting a fraction of a second.
(2) The encoder breaks each frame down into 72 discrete frequency bands and analyzes the audio signal to determine its 'spectral energy distribution'. It looks for parts of the frequency spectrum that have a lot of sound energy in them and parts that have almost none. Basically, the algorithm tries to figure out where the most important frequencies are in the sound.
(3) The encoder then decides how much data to retain and how much to discard, depending on the size of the mp3 that the user wants as an output file. Relative size and rate of data compression for mp3 files are measured in kilobytes per second, this is because the size of an mp3 is a measure of bandwidth for many applications – the whole point of encoding is miniaturization. So, a bigger mp3 file has more kilobytes per second. To make a smaller mp3, the encoder has to throw away more data from the original CD recording: it will discard more data when it makes a 64kbps file than a 128kbps file.
(4) The encoder calculates a new timbral measurement for each frame based on what it learned about the shape of the incoming signal and on a mathematical table of values that represents human psychoacoustic response. The point is to get rid of data that people cannot hear.
(5) The encoder then runs through Huffman coding, which is a standard data compression algorithm designed to eliminate most redundant data in the file. Similar to a zip file, it does not get rid of any data per se, but spatially consolidates data storage.
(6) Finally, the encoder assembles a 'serial bitstream' which contains header information and instructions for each frame. These instructions are for playback programs and devices to ensure consistent playback.

As a form of data compression, the most compelling part of the mp3 is the psychoacoustic model encoded within it. To personify the technology, it presumes that the sense of hearing discards most of the sound that it encounters, attempting to imitate the process by which the human body discards soundwaves in the process of perception. It preemptively discards data in the soundfile that it anticipates the body will discard later, resulting in a smaller file.

That the ear (or any sense) acts as a filter is an old idea. From Aristotle on the senses to Aldous Huxley on psychedelic drugs, it is well established that the senses do not mimetically reproduce the world they encounter – they shape it (Aristotle, 1976[350BC]; Huxley, 1954). Today, the explanation is a bit more clinical. Put simply, the auditory nerve fires with less frequency than the frequencies of sound. The nerve in the inner ear cannot keep up with sound as it actually happens. Yet somehow, between the cochlea, the auditory nerve and the auditory center in the brain, people get a sense of the detailed rise and fall of sounds. Scholars of psychoacoustics have proposed a number of analyses as to why the ear works in the way that it does, but no one theory is dominant (Mathews, 2001a). The key point is that while traditionally, sound reproduction technologies have been theorized in terms of their relation of absolute fidelity to a sound source, the human ear is not capable of such fine distinctions. In fact, people can lose most of the vibrations in a recorded sound and still hear it as roughly the same sound as the version with no data compression. This is the principle upon which the mp3 rests.

This discussion of psychoacoustics is not meant to be 'psychological' as humanities scholars generally understand the term. Psychoacoustics is a crucial component of the embodiment of sound. Psychoacoustic response is not localized in a single place: the ear concentrates, focuses and stratifies vibrations into sounds, which the auditory nerve translates into signals for the brain to perceive. But the ear canal is not the only place that conducts vibrations to the auditory nerve: the whole head can and does conduct vibration, as does the chest cavity. For example, scholars of psychoacoustics have devised frequency-response curves for different body parts: the ear canal and drum, the 'sphere' of the head, the pinna (outer ear), the torso and the neck (Mathews, 2001b). The jaw, in particular, is a useful conduit of sound. Some early hearing aids were held in the teeth rather than placed in the ear; the notoriously hard-of-hearing Thomas Edison's bitemarks can be found on many early phonograph prototypes. So what we have here is a psychacoustic body, a body that 'does stuff' to vibration in order to turn it into sound. Sound is a product of perception, not a thing 'out there' – the only thing 'out there' is vibration, which the body organizes and stratifies into what we call sound (Sterne, 2003). The body shapes vibrations before they enter the ear and become sound. This is obvious in cases of tinnitus (ringing of the ears) and frequency-dependent hearing loss. But it is also true of people who have undamaged hearing. Psychoacoustic effects are based on the fact that the body creates sound by organizing vibrations.

Mp3s use psychoacoustic principles to get rid of the sounds that we supposedly would not hear anyway. There are three specific psychoacoustic tricks that mp3 encoders use to reduce the size of data files: simultaneous or auditory masking, temporal masking and spatialization. Auditory masking is

the elimination of similar frequencies, based on the principle that when two sounds of similar frequency are played together and one is significantly quieter, people will hear only the louder sound. Temporal masking is a similar principle across time: if there are two sounds very close together in time (less than about five milliseconds apart, depending on the material) and one is significantly louder than the other, listeners can only hear the louder sound. The third principle is spatialization. While it is very easy to locate the direction of sounds in the middle of the audible range when they are played back in stereo, it is close to impossible for people to locate very low or very high sounds. To save more dataspace, the mp3 encoder saves sounds at either end of the frequency spectrum only once for both channels, rather than twice and plays them back as mono files. Since most human adults cannot hear above 16khz, some mp3 encoders also throw out all the data from 16–20khz to save even more space. Psychoacoustically, the mp3 is designed to throw away sonic material that listeners supposedly would not hear otherwise.

This process requires a good deal of compromise. Mp3s of songs do not sound the same as the CD recordings; a professional audio engineer could certainly tell the difference. But the amazing thing is that as we move from ideal listening environments into the situations in which people usually hear mp3s, it becomes increasingly difficult to distinguish. Mp3s are designed to be heard via headphones while outdoors, in a noisy dorm room, in an office with a loud computer fan, in the background as other activities are taking place and through low-fi or mid-fi computer speakers. They are meant for casual listening, moments when listeners may or may not attend directly to the music – and are therefore even *less* likely to attend to the sound of the music. In other words, the mp3 is a medium which, in most practical contexts, gives the full experience of listening to a recording while only offering a fraction of the information and allowing listeners' bodies to do the rest of the work. The mp3 plays its listener. Built into every mp3 is an attempt to mimic and, to some degree preempt, the embodied and unconscious dimensions of human perception in the noisy, mixed-media environments of everyday life.

It has been suggested already here that the portability of mp3s was built into their form by an industry which believed that exchange and compatibility were in their objective interests. Some of the innovations were technical, but a whole other set of innovations dealt with how people hear things. Since the 1910s, AT&T's Bell Laboratories has researched ways to cram more sound into limited bandwidth telephone lines because it would allow a massive increase in signal exchange with no expansion of infrastructure.[7] For example, although you may think that you hear my booming voice on the other end of the telephone line, you actually only hear what Hermann Helmholtz (1954[1863]) called the 'upper partials' of

the signal. Helmholtz showed that when certain higher frequencies are played together, they effectively synthesize a sound an octave lower. The phone system does not have to transmit any low-frequency signal – your ears and brain will simply fill it in.

The audio media that we encounter every day use a whole set of techniques such as this to shape our sonic environments. But the portability also exists in the psychoacoustic dimensions. The mp3 fetishizes and makes use of the imperfections of healthy hearing while presuming a so-called normal listening situation. The ideal listener implied by the mp3's psychoacoustic coding is Theodor Adorno's nightmare: the 'distracted' consumer of mass culture (2002[1938], 1993[1945]). In a media-saturated environment, portability and ease of acquisition trumps monomaniacal attention. Of course, actual listening practices and environments vary widely, and as Michael Bull (2000) has shown with portable stereos, the meanings attributed to a sound technology by its users are quite variable depending on its actual use. But the point to take here is simple: at the psychoacoustic level as well as the industrial level, the mp3 is designed for promiscuity. This has been a long-term goal in the design of sound reproduction technologies.

AROUND THE MP3

Although it stands in a long line of technologies that have made use of acoustic and psychoacoustic principles, the mp3 has applied psychoacoustics to a much greater degree than any major sound reproduction technology that preceded it. The embodied mp3 stands in stark contrast to the concept of 'the virtual', which has received so much play in cultural analyses of digital media. Countless scholars have treated virtuality as an ontological dimension of digital media or their necessary consequence. Virtuality is supposed to separate the subject from the body – and digital media are supposed to be the most radical form of mediated disembodiment yet invented. Ken Hillis (1999) has shown the very embodied dimension of virtual reality. After all, the typical virtual reality set has both a pair of goggles and glove – it is both visual and haptic. The mp3 suggests an even more radical challenge to the concept of virtuality because of its direct and sensuous interaction with an embodied, sensing, unthinking subject. If sound is not 'out there' but rather created by the process of perception, then the mp3 is not a simulation of sound or a virtual sound. It is simply another mode through which the effect of sound is produced and embodiment is the defining characteristic of the experience. The subject of the mp3 is almost the opposite of the supposed intending, self-knowing, consciously self-constructing subject of virtual reality.

The history of digital audio has been read largely by cultural scholars as being about the relationship of originals and copies and especially about questions of the fidelity of copies to 'original sounds' that exist outside the

process of reproduction (Corbett, 1994; Mowitt, 1987; Rothenbuhler and Peters, 1997). This philosophical position has come under some criticism because it separates the ontology of reproduced sound from the social situation of sound reproduction (Lastra, 2000). But even on so-called ontological terms, there are questions to be raised: the point of recording something is not simply to reproduce the event later on, but also to move the recording across space. Recording has both space and time-binding characteristics (Innis, 1951). The portability of recordings is as important a feature of their history as the nature of their reproduction. And it is upon the terrain of portability, rather than fidelity, that we encounter the mp3.

One could say that mp3s are like other audio container technologies because they distort or change the sound as they reproduce it. Distortion is truth in the world of recorded audio (Poss, 1998). The point of recording and reproduction is not to mirror sound but to shape it actively. One of the important shifts in the history of sound reproduction was from an imitation of the causes of sound to imitations of sound as an effect. Rather than attempting to reproduce the mechanism of a violin or a voice, telephones, phonographs, radios, microphones and their cousins reproduced the vibrations of a diaphragm modeled on the tympanic membrane in the human middle ear. Anything containing devices in it to change sound to signal or signal to sound – from stereo speakers to cellphones – is a tympanic technology (Sterne, 2003). The mp3 is a supplement to tympanic sound reproduction. Not only is the vibration of the diaphragm a process that can be imitated and induced as a result of the imitation (your eardrums vibrate because the speakers in the room do), but it models, and to some degree imitates, the process of sound perception. Bruno Latour (1988) calls this aspect of technological practice 'delegation', where people delegate a function that they once performed to a machine. His classic example is the door-closer, which closes doors for people. Mp3s and their encoders simply extend this logic. Tympanic technologies are machines designed to hear for people and mp3s are designed to perceive for them.

Mp3 technology also has an interesting relationship with other bodily technologies of communication. The mp3 works automatically on the body. Mp3 listening might involve 'practical knowledge' (Bourdieu, 1990), where the body goes through routines that do not enter the conscious mind. Certainly, mp3 listening requires a whole set of bodily techniques, dispositions and attitudes. But the mp3 goes even further than this. The encoded mathematical table inside the mp3 that represents psychoacoustic response suggests less a 'technique of the body' as these authors would have it, than a concordance of signals among computers, electrical components and auditory nerves.

Scale matters here: quantitatively speaking, the coordinated movements of mp3 sonics are so much smaller than the movements produced by a

socialized body, that we may be talking about a *qualitative* difference between listening practice as a technique of the body and the mp3 as a concordance of signals. In the techniques of the body, bodily movement is conditioned as part of the socialization process. People learn how to walk, sit, gesture, etc. Certainly, there is a whole set of techniques of listening that are assumed or enabled by mp3 technology. But mp3s also *anticipate* the tiny movements of the inner ear, which are not so much organized and disciplined in the mathematical table as they are modeled and anticipated. Thus, the mp3 uses a construct of the body to modify data, electrical signals and eventually sounds before they get to listeners' ears. This is to say that mp3s require body techniques, but are not themselves body techniques in the same sense – they model perception in order to affect it.

The mp3 carries with it a sonic logic of resourcing and supply. At the risk of overextending a metaphor, the logic could be cast in economic terms: the mp3 encoding process puts the body on a sonic austerity program. It decides for its listeners what they need to hear and gives them only that. Listeners' bodies, brains and ears then contribute a kind of surplus activity (if not quite labor) to make the system run. Of course, commentators from *Wired Magazine* to Courtney Love have touted the mp3 as some kind of sonic liberation. But it represents a liberation of just-in-time sound production, where systems give listeners less and ask their bodies to do more of the work.

Despite the attractiveness of the economic metaphor, it does not quite work when it comes to assessing the mp3 as a cultural artifact. The mp3 is not nearly as nefarious as neoliberal economic policies. One could rightly argue that rather than being a cruel exploitation of the limits of auditory perception, the mp3 encoder instrumentalizes and even celebrates the limits of the human ear. It suggests how little 'input' people need in order to have powerful and significant aesthetic experiences. Perhaps the best attitude, then, is a certain ambivalence toward mp3s. True, people do not get the 'full' sonic experience of recordings when listening to them in mp3 format. But given the infrequency with which people are in a position to have the 'full' sonic experience of recorded music, perhaps the trade-off is worth it in some cases. At the very least, the success of the mp3 adds a new twist to the critique of correspondence theories of representation and their corresponding aesthetics. To use a term from Ivan Illich (1973), the mp3 is a comparably 'convivial' technology for listening to mediated music in noisy, multimedia or otherwise distracting (or 'distracted') contexts. It is all the more remarkable because it grew up amid such unconvivial technical systems as the recording industry, recording studios, CD plants and computer networks.

The mp3 is a form designed for massive exchange, casual listening and massive accumulation. As a container technology designed to execute a

process on its contents, it does what it was made to do. The primary, illegal uses of the mp3 are not aberrant uses or an error in the technology; they are its highest moral calling: 'Eliminate redundancies! Reduce bandwidth use! Travel great distances frequently and with little effort! Accumulate on the hard drives of the middle class! Address a distracted listening subject!' These are the instructions encoded into the very form of the mp3. This is the mission that an mp3 carries out as it travels down network lines onto my hard drive; as it instructs my computer to construct a datastream that will become electricity, that in turn will vibrate the speakers on my desk and the membranes in my ears as I type this sentence. The mp3 has a job to do, and it does it very well.

Acknowledgements

Many thanks to Carrie Rentschler, Emily Raine, Jeremy Morris, Will Straw, Jenny Burman, Louise Meintjes, Larry Grossberg, Ken Hillis, Ken Wissoker, Lisa Nakamura, Christian Sandvig, Alex Mawyer, Brian Horne, two anonymous reviewers and audiences in Montreal, Durham, Madison, Pittsburgh and Chicago for their comments on earlier versions of this article.

Notes

1 *Pace* Lev Manovich (2001), who has argued that 'new media' is a more robust phrase than 'digital media' or some other name, most so-called 'new' media are not that new anymore. The term 'new' is incredibly value-laden in our heavily commercialized culture, but I use it here and throughout the article with the recognition that 'new media' signifies a fairly coherent set of objects of study and a number of emergent intellectual traditions.

2 These three categories indicate three different kinds of container technology for Sofia. Parsing them out is not necessary for the current argument, but interested readers should refer to her article.

3 I use the term 'data compression' to signal the difference between the processes that remove data from audio files (traditionally called 'compression', which only can be accomplished digitally) and the process of reducing the distance between the loudest and quietest points in an audio signal. This is also called 'compression', but it can be accomplished through either analog or digital signal processing.

4 The ISO is a network of national standards institutes from 148 countries which collaborate with international organizations, governments, industry, business and consumer representatives. The International Electrotechnical Commission focuses on standards for electronic and magnetic devices and is now affiliated with the World Trade Organization.

5 A full academic history of the mp3 has yet to be written. The only thorough journalistic history currently available is Bruce Haring's *Beyond the Charts* (2000). Taking his cues from intellectual property debates, he presents the mp3 as part of a longer story about digital audio, online distribution and the music industry.

6 Forthcoming work will explore this question in greater detail.

7 The line of psychoacoustic research first developed at Bell Laboratories in the 1910s is the genealogical point of origin for digital audio as we know it today. It is also an important origin point for cybernetic models of communication.

References

Adorno, T. (1993[1945]) 'A Social Critique of Radio Music', in N. Strauss (ed.) *Radiotext(e)*, pp. 272–9. New York: Semiotext(e).

Adorno, T. (2002[1938]) 'On the Fetish Character of Music and the Regression of Listening', in R. Leppert (ed.) *Adorno: Essays on Music*, pp. 288–317. Berkeley, CA: University of California Press.

Apple (2004) 'iPod Mini', *Wired* 12(March), back cover.

Aristotle (1976[350BC]) *De Anima* (trans R.D. Hicks). New York: Arno Press.

Atkinson, J. (1999) 'Mp3 and the Marginalization of High End Audio', *Stereophile* (February): 22.

Attali, J. (1985) *Noise: the Political Economy of Music* (trans B. Massumi). Minneapolis, MN: University of Minnesota Press.

Barbrook, R. (1998) 'The Hi-Tech Gift Economy', *First Monday* 3(12) (available at: http://www.firstmonday.org/issues/issue3_12/barbrook/index.html).

Benjamin, W. (1968[1936]) *Illuminations* (trans. H. Zohn). New York: Schocken.

Bettig, R. (1997) 'The Enclosure of Cyberspace', *Critical Studies in Mass Communication* 14(2): 138–58.

Bijker, W. (1995) *Of Bicycles, Bakelites and Bulbs: Toward a Thoery of Sociotechnical Change*. Cambridge, MA: MIT Press.

Bourdieu, P. (1990) *The Logic of Practice* (trans. R. Nice). Stanford, CA: Stanford University Press.

Bull, M. (2000) *Sounding Out the City: Personal Stereos and Everyday Life*. New York: New York University Press.

Burkart, P. and T. McCourt (2004) 'Infrastructure for the Celestial Jukebox', *Popular Music* 23(3): 349–62.

Corbett, J. (1994) *Extended Play: Sounding Off from John Cage to Dr. Funkenstein*. Durham, NC: Duke University Press.

Eide, O. (2001) 'Bob Ludwig', *Mix*, 1 December (available at: http://mixonline.com/mag/audio_bob_ludwig/index.html).

Evangelista, B. (2003) '4 Students Sued Over Music Trading Software; Record Industry Goes After Campus File-Sharing Programs', *San Francisco Chronicle,* 4 April: 2.

Hacker, S. (2000) *Mp3: The Definitive Guide*. Cambridge, MA: O'Reilly.

Hall, S. (1992) 'Cultural Studies and its Theoretical Legacies', in L. Grossberg, C. Nelson and P. Treichler (eds) *Cultural Studies*, pp. 277–94. New York: Routledge.

Haring, B. (2000) *Beyond the Charts: Mp3 and the Digital Music Revolution*. Los Angeles, CA: OTC Books.

Helmholtz, H. (1954[1863]) *On the Sensations of Tone as a Physiological Basis for the Theory of Music* (2nd edn, trans. A.J. Ellis). New York: Dover.

Hillis, K. (1999) *Digital Sensations: Space, Identity and Embodiment in Virtual Reality*. Minneapolis, MN: University of Minnesota Press.

Huxley, A. (1954) *The Doors of Perception*. New York: Harper.

Illich, I. (1973) *Tools for Conviviality*. New York: Harper & Row.

Innis, H. (1951) *The Bias of Communication*. Toronto: University of Toronto Press.

Jones, S. (2000) 'Music and the Internet', *Popular Music* 19(2): 217–30.

Jones, S. (2002) 'Music that Moves: Popular Music, Distribution and Network Technologies', *Cultural Studies* 16(2): 213–32.

Krim, J. (2003) 'File Sharing Forfeits Right to Privacy: Judge Tells Verizon to Identify Customer', *Washington Post,* 25 April: 1.

Lastra, J. (2000) *Sound Technology and American Cinema: Perception, Representation, Modernity*. New York: Columbia University Press.

Latour, B. (1988) 'Mixing Humans and Nonhumans Together: The Sociology of a Door-Closer', *Social Problems* 35(1): 298–310.
Latour, B. (1996) *Aramis, Or the Love of Technology*. Cambridge, MA: Harvard University Press.
Lazin, D. (2003) 'New Levy Boosts Price of Mp3 Players', *Edmonton Journal,* 13 December: 1.
Lessig, L. (2000) *Code and Other Laws of Cyberspace*. New York: Basic Books.
Lessig, L. (2002) *The Future of Ideas: The Fate of the Commons in a Connected World*. New York: Vintage Books.
Locke, J. (1692) *Some Considerations on the Consequences of the Lowering of Interest and Raising the Value of Money*. London: A. & J. Churchill.
McCollum, K. (2000) 'Students Jam College Networks with Use of Napster to Find Audio Files', *Chronicle of Higher Education* 25, 25 February: A50.
McCourt, T. and P. Burkart (2003) 'When Creators, Corporations and Consumers Collide: Napster and the Development of On-Line Music Distribution', *Media, Culture & Society* 25(3): 333–50.
McLeod, K. (2001) *Owning Culture*. New York: Peter Lang.
McLeod, K. (2005) *Freedom of Expression: Tales from the Dark Side of Intellectual Property Law*. New York: Doubleday.
Manovich, L. (2001) *The Language of New Media*. Cambridge, MA: MIT Press.
Marx, K. (1967[1867]) *Capital: A Critical Analysis of Capitalist Production*, Vol. 1 (trans. S.M.A.E. Aveling). New York: International Publishers.
Mathews, M. (2001a) 'The Auditory Brain', in P.R. Cook (ed.) *Music, Cognition and Computerized Sound: An Introduction to Psychoacoustics*, pp. 11–20. Cambridge, MA: MIT Press.
Mathews, M. (2001b) 'The Ear and How It Works', in P.R. Cook (ed.) *Music, Cognition and Computerized Sound*, pp. 1–10. Cambridge, MA: MIT Press.
Mitchell, J.L., W.B. Pennebaker, C.E. Fogg and D.J. LeGall (eds) (1997) *MPEG Video Compression Standard*. New York: Chapman & Hall.
Mowitt, J. (1987) 'The Sound of Music in the Era of its Electronic Reproducibility', in R. Leppert and S. McClary (eds) *Music and Society: the Politics of Composition, Performance and Reception*, pp. 173–97. New York: Cambridge University Press.
Mumford, L. (1959) 'An Appraisal of Lewis Mumford's Technics and Civilization', *Daedalus* 88: 527–36.
Mumford, L. (1966) *The Myth of the Machine: Technics and Human Development*, Vol. 1. New York: Harcourt, Brace and World.
Pinch, T. and W. Bijker (1984) 'The Social Construction of Facts and Artefacts: Or How the Sociology of Science and the Sociology of Technology Might Benefit Each Other', *Social Studies of Science* 14(3): 399–441.
Poss, R. (1998) 'Distortion Is Truth', *Leonardo Music Journal* 8: 45–8.
Regnier, P. (2003) 'Digital Music for Grown-Ups', *Money*, July: 113.
Rothenbuhler, E. and J.D. Peters (1997) 'Defining Phonography: An Experiment in Theory', *Musical Quarterly* 81(2): 242–64.
Sherburne, P. (2003) 'Digital DJing App that Pulls You In', *Grooves* 10: 46–7.
Slack, J.D. (1984) *Communication Technologies and Society: Conceptions of Causality and the Politics of Technological Intervention*. Norwood, NJ: Ablex.
Smith, A. (1993[1776]) *An Inquiry into the Nature and Causes of the Wealth of Nations*. New York: Oxford University Press.
Sofia, Z. (2000) 'Container Technologies', *Hypatia* 15(2): 181–219.

Stabile, C. (1994) *Feminism and the Technological Fix*. New York: Manchester University Press.
Sterne, J. (2003) *The Audible Past: Cultural Origins of Sound Reproduction*. Durham, NC: Duke University Press.
Winner, L. (1986) *The Whale and the Reactor: A Search for Limits in the Age of High Technology*. Chicago, IL: University of Chicago Press.
Wise, J.M. (1997) *Exploring Technology and Social Space*. Thousand Oaks, CA: Sage.
Wood, C. and D.A. Jenish (2000) 'Free Music!', *Maclean's,* 20 March: 42.

JONATHAN STERNE teaches in the Department of Art History and Communication Studies and the History and Philosophy of Science Program at McGill University, Montreal. He is author of *The Audible Past: Cultural Origins of Sound Reproduction* (Duke University Press, 2003) and numerous essays on media, technologies and the politics of culture. His next book is tentatively titled *MP3: The Meaning of a Format.* He is also an editor of *Bad Subjects: Political Education for Everyday Life* (www.badsubjects.org), one of the longest continuously-running publications on the internet. *Address*: Department of Art History and Communication Studies, McGill University, Arts W-225, 853 Sherbrooke St W., Montreal QC, H3A 2T6, Canada. [email: jonathan.sterne@mcgill.ca]